

Caracterização de genes através de recursos bioinformáticos

Lurdes Jorge e Madalena Vaz

ESA-IPB, CIMO

Após descodificação da fase de leitura aberta de um gene, uma série de ferramentas bioinformáticas podem ser utilizadas para a caracterização da sequência deduzida da proteína.

Uma pesquisa no website do ExPASy Proteomics Server (<http://expasy.org/tools>) e uma sequência nucleotídica permitem-nos identificar e caracterizar proteínas, identificar motivos, padrões e perfis, inferir a sua estabilidade, localização celular ou função, fazer as predições das estruturas secundária e terciária, procurar sequências similares depositadas em bases de dados e compará-las, estabelecer relações filogenéticas.

Neste caso usámos a ORF ("open reading frame") do gene *lip2* de *Trichoderma harzianum*, cuja sequência pode ser acedida na base de dados da EMBL com o número de acesso [AM774154.1](#).

A proteína codificada por *lip2* tem 404 aminoácidos, massa molecular calculada de 44604,3 Dalton e ponto isoeléctrico global calculado de 5,0. É considerada estável.

Uma pesquisa efectuada na base integrada InterProScan incluiu *lip2* na família das lipases com serina no centro activo (PROSITE PS00120, posição 210-219), e na das lipases_classe3 (Pfam PF01746, posição 131-190) com uma probabilidade de $1,4.e^{-31}$.
<http://www.ebi.ac.uk/Tools/pfa/iprscan/>

A sua estrutura primária está de acordo com o consenso **G-x-S-x-G**, descrito como centro activo de lipases, inserido na sequência consenso de lipases com serina no centro activo: **[LIV]-{KG}-[LIVFY]-[LIVMST]-G-[HYWV]-S-{YAG}-G-[GSTAC]**.

Os primeiros 25 aminoácidos constituem uma sequência sinal, que sugere a entrada desta proteína no retículo endoplasmático, e uma localização extracelular.

No que respeita a modificações pós-traducionais, entre outras, a proteína codificada por *lip2* apresenta três possíveis sítios de N-glicosilação, quatro sítios prováveis de N-miristoilação, um de palmitoilação e 28 sítios potenciais de fosforilação. A palmitoilação aumenta a superfície hidrofóbica e a afinidade para os substratos e desempenha um papel importante no transporte de proteína.

O sistema de redes neurais do "Pôle BioInformatique Lionnais/Network Protein Sequence Analysis" (<http://npsa-pbil.ibcp.fr>) prediz para a proteína lip2 uma estrutura secundária típica das α/β hidrolases, em que há alternância de domínios de conformação em hélice α com domínios de conformação em folha β .

A comparação da sequência nucleotídica de *lip2* com o banco de sequências de proteínas (UNIPROT) pôs em evidência a homologia existente entre lip2 e outras lipases fúngicas, nomeadamente as de alguns fungos fitopatogénicos importantes. Os resultados foram obtidos por utilização do algoritmo de FASTA. (<http://www.ebi.ac.uk/Tools/sss/fasta/>).